Uncertainty quantification of fast sodium current steady-state inactivation for multiscale models of cardiac electrophysiology: **supplemental material**

Pras Pathmanathan, Matthew S. Shotwell, David J. Gavaghan,
Jonathan M. Cordeiro and Richard A. Gray

S1 Statistical analysis

Outlier assessment

Figure S1 repeats Figure 4 (top left) using all records in the *Cord-Epi* dataset. The estimate of $\mathbf{p}_I$ corresponding to the visually anomalous experimental result lies well outside the 95% prediction region (computed using the complete data), and therefore we conclude that the experimental curve is statistically an outlier and that its exclusion is warranted.

Model diagnostics

Figure S2 plots the inactivation residuals (difference between observed and fitted inactivation value), against the fitted value, for the NLME fits to each of the four datasets. The underlying assumptions used by this method imply that these residuals should be normally-distributed with zero mean and constant variance ('homoscedastic'). This assumption cannot hold at very small or high voltages (when the fitted value is near 0 or 1), since the function being fit exponentially nears 0 or 1, and yet experimental observations are necessarily ≥ 0 and ≤ 1. However, this modelling inconsistency does not significantly affect the inferences that arise under the NLME method, as applied to these data. We see that, except for near fitted value equal to 0 or 1, the residuals do not exhibit any broad nonlinear trends, and are roughly homoscedastic. Figure S3 provides the corresponding 'quantile-quantile' plots to evaluate the residual normality assumption. The quantile-quantile plots do not exhibit marked deviation from a linear trend. Hence, the normality assumption is supported.

As a further test of robustness to parametric distributional assumptions, we implemented a nonparametric bootstrap alternative to construct 95% confidence regions associated with $\mathbf{p}^{\mathrm{NLS}}$ and $\mathbf{p}^{\mathrm{NLME}}$. The nonparametric bootstrap is robust to a variety of deviations from residual normality, heteroscedasticity, and other conditions that violate the parametric assumptions that underpin the NLME method. The bootstrap confidence regions, plotted in Figure S4, are somewhat less conservative than the Wald-type regions associated with $\mathbf{p}^{\mathrm{NLME}}$, but are similar in size and orientation. The boostrap regions associated with the weighted NLS estimates are markedly different to the corresponding Wald-type confidence regions, raising additional concern regarding the weighted NLS approach.

Overall, we conclude that the function $F_2(V; V_0, K)$ fits the data well, and that the NLME framework provides adequate flexibility in modelling population variability.

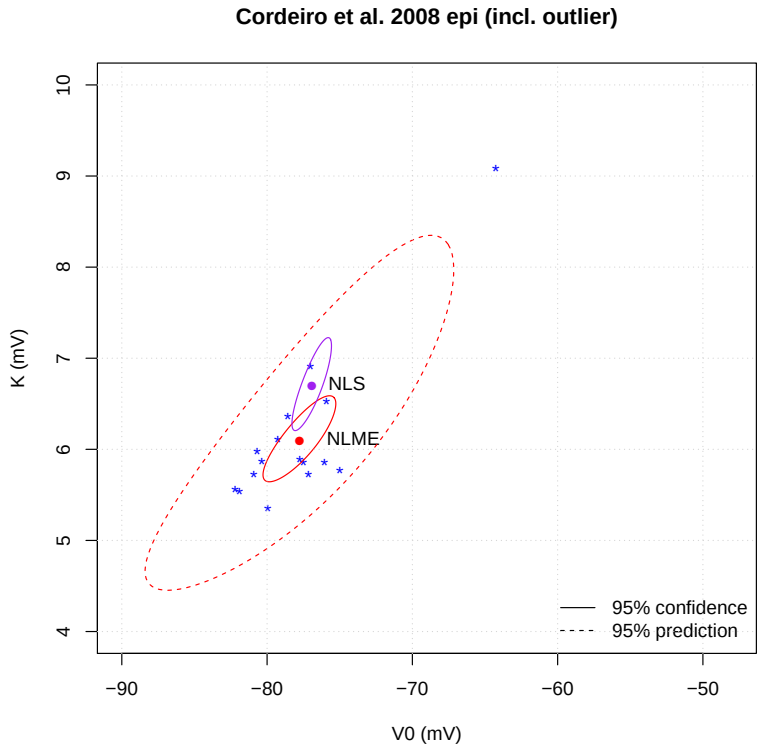

Figure S1: Figure 4 (top left) repeated without removal of potential outlier.

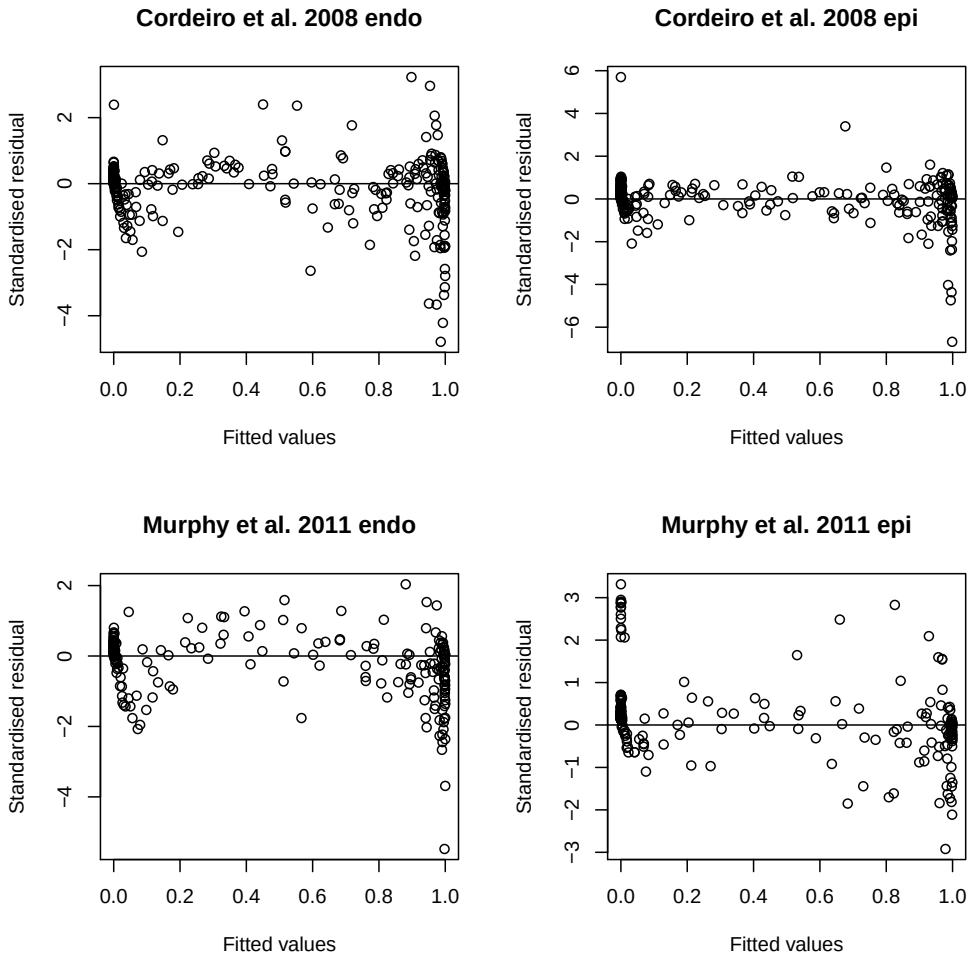

Figure S2: Residuals (difference between observation and fit), against the fitted value, for the NLME fits for each of the four datasets.

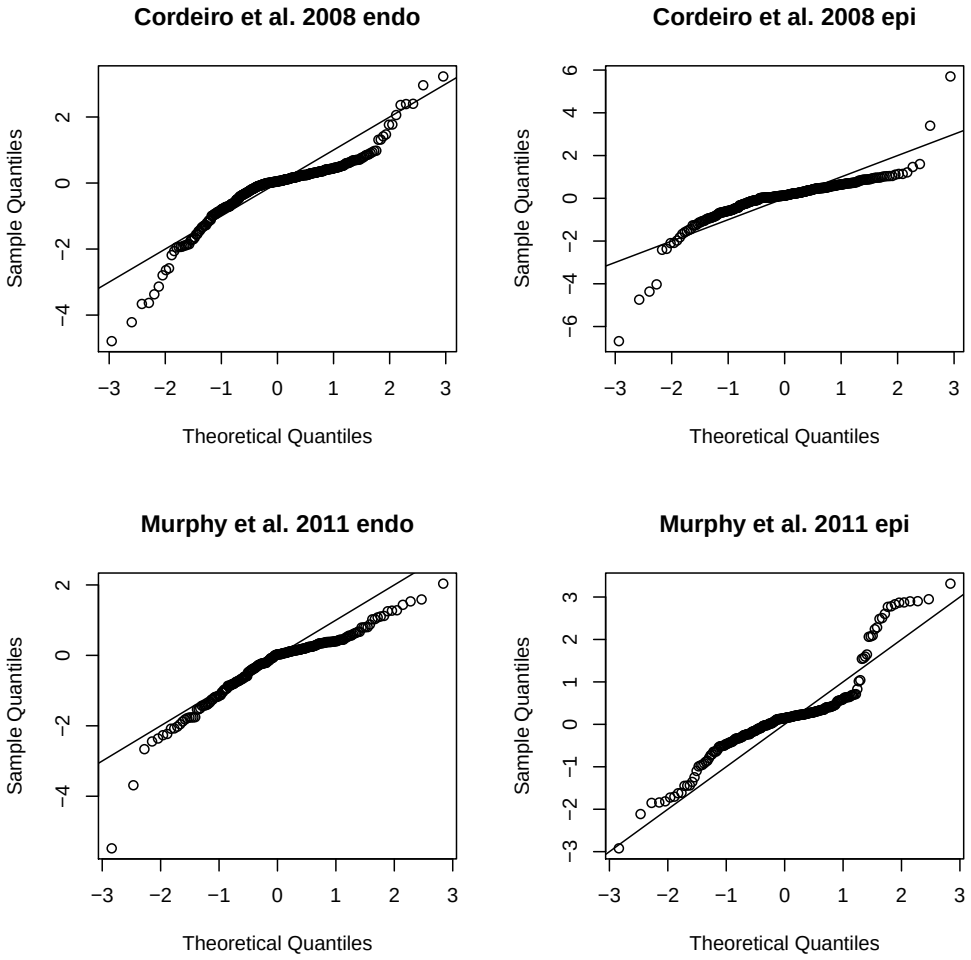

Figure S3: Quantile-quantile plot for the NLME fits with each of the four datasets.

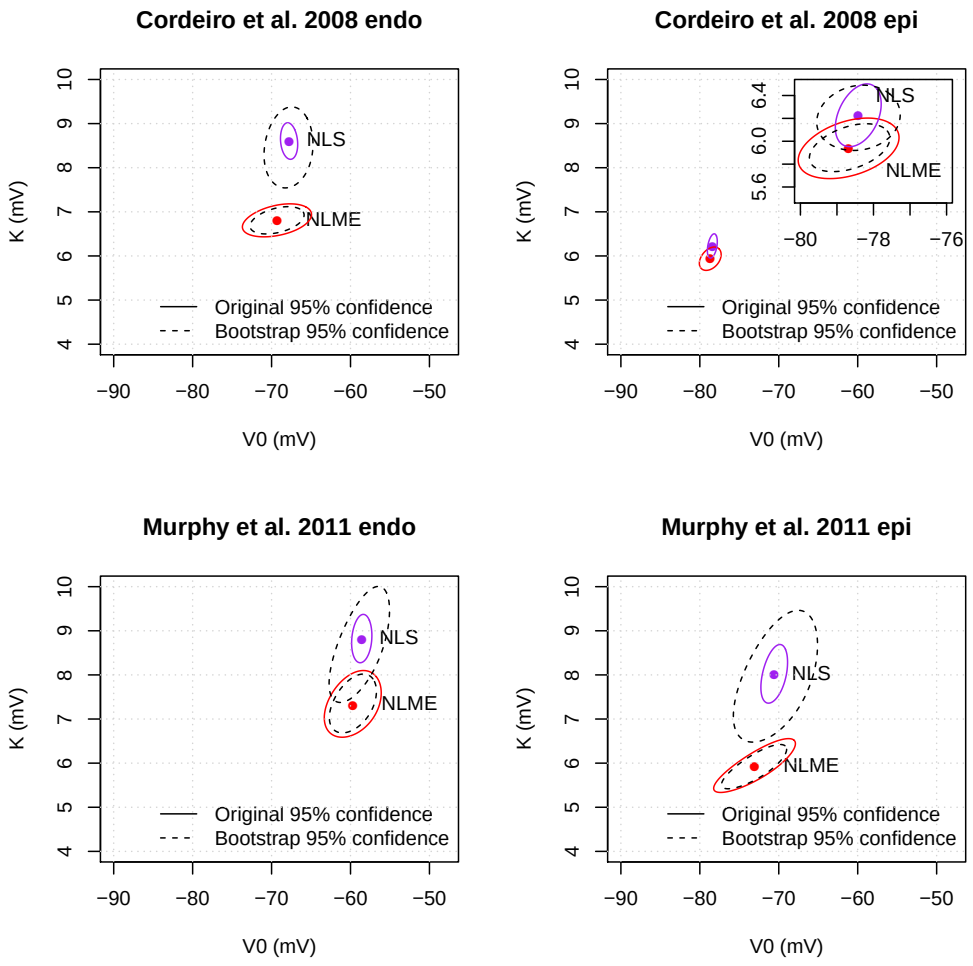

Figure S4: 95% Confidence regions computed using the nonparametric bootstrap method, together with NLME and NLS confidence regions.

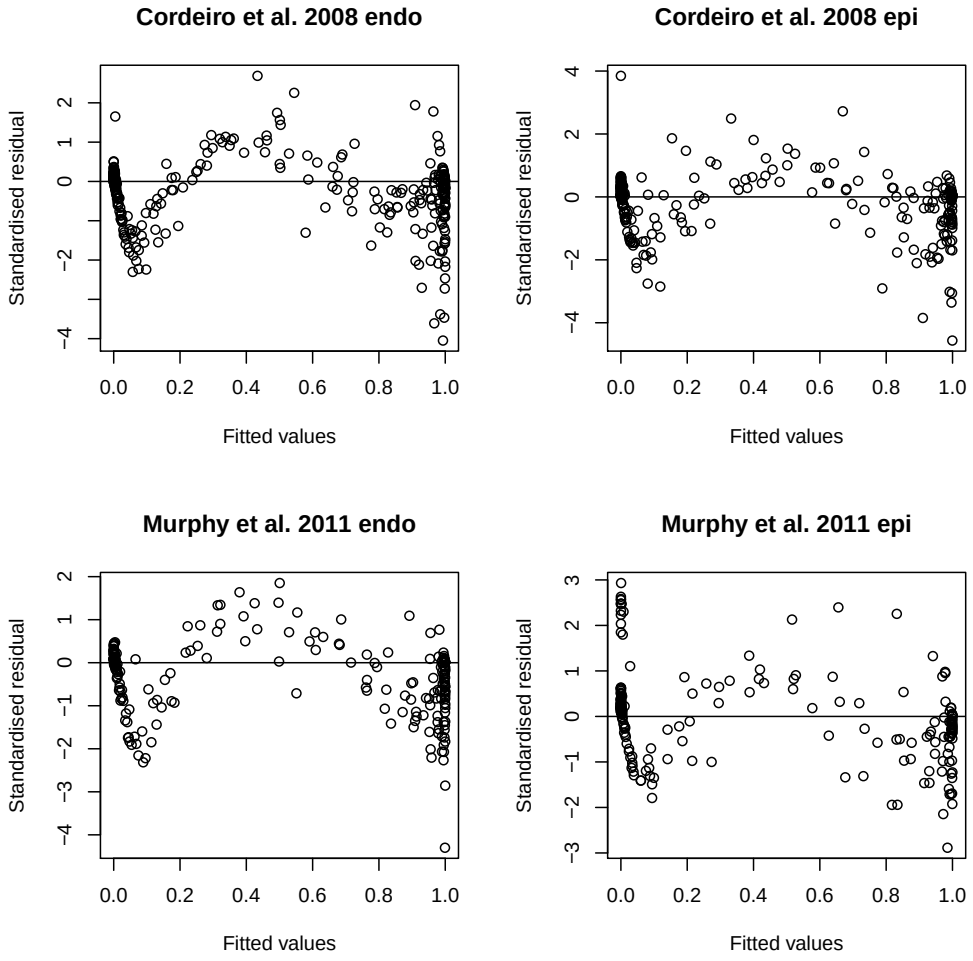

Figure S5: Residuals (difference between observation and fit), against the fitted value, when the function $F_1(V; V_0, K)$ is used instead of $F_2(V; V_0, K)$.

Choice of n

Figure S5 plots the residuals against fitted values if the function $F_1(V; V_0, K)$ is instead fitted to the data. Clear nonlinear trends are observed, suggesting that $F \equiv F_2$ provides a better fit than F_1.

Generalised nonlinear least squares

The nonlinear least squares method may be further generalised by specifying a correlation structure among observations. For example, in the current context, deviations from the average inactivation are similar among observations from an individual cell, and such deviations are most similar in consecutive measurements (i.e., consecutive voltage levels). The so-called first-order autoregressive (AR[1]) model accounts for such correlation.

In order to partially visualise the distinctions between the weighted NLS, generalised NLS

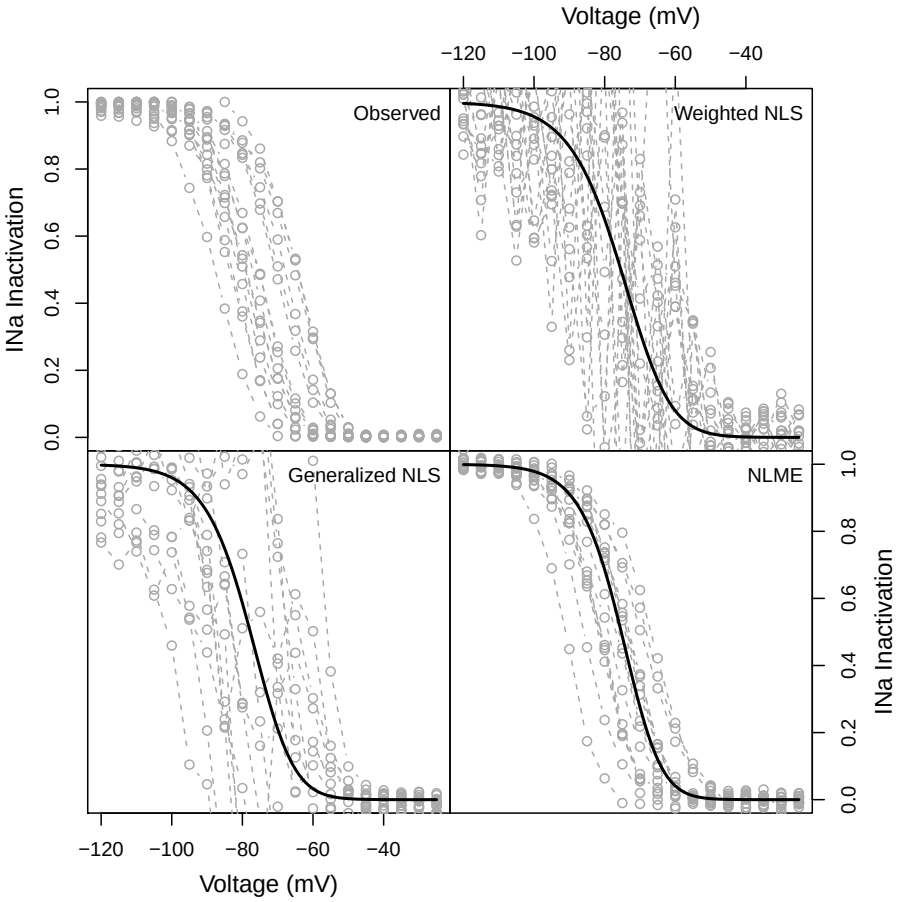

Figure S6: In clockwise order, observed and simulated I_{Na} inactivation data using the weighted NLS, generalised NLS, and NLME methods (dashed curves), with overlaid fits corresponding to $\mathbf{p}^{\text{NLS}}$, generalised $\mathbf{p}^{\text{NLS}}$, and $\mathbf{p}^{\text{NLME}}$ (solid curves), respectively.

(i.e., with AR[1] covariance model), and NLME methods, we note that each method encodes a probabalistic model for the mechanism that gave rise to the experimental data. As such, we may use the fitted models to simulate new data that might arise under an identical experimental design.

Figure S6 illustrates the observed *Cord-Endo* data, and newly simulated data using the weighted NLS, generalised NLS (AR[1] covariance model), and NLME methods. Although the generalised NLS simulations offer an improvement to those associated with the weighted NLS model, the NLME simulations are clearly most similar to the observed data.

S2 Uncertainty propagation

Computation of probability distributions

In this section we describe how the probability distributions for upstroke velocity were computed for Figure 6.

We use a simple numerical integration approach based a 100×100 grid over parameter space. This grid is the same as that used to plot the upstroke velocity in Figure 6 (top left). Let $y \equiv y(\mathbf{p})$ be the quantity of interest (i.e. upstroke velocity), dependent on uncertain input $\mathbf{p} = (V_0, K)$, and let Y be the corresponding random variable. Then the probability that Y lies between values Y_1 and Y_2 is

$$\mathbb{P}(Y \in (Y_1, Y_2)) = \int_{\{\mathbf{p} \,:\, y(\mathbf{p}) \in (Y_1, Y_2)\}} f_{\mathbf{p}}(\mathbf{p}) \, \mathrm{d}p_1 \mathrm{d}p_2,$$

where $f_{\mathbf{p}}$ is the probability density function for $\mathbf{p}$. Alternatively, treating y as a function of $\mathbf{q} = (V_0, \log(K))$ (see Section 2.1.2),

$$\mathbb{P}(Y \in (Y_1, Y_2)) = \int_{\{\mathbf{q} \,:\, y(\mathbf{q}) \in (Y_1, Y_2)\}} f_{\mathbf{q}}(\mathbf{q}) \, \mathrm{d}q_1 \mathrm{d}q_2,$$

where $f_{\mathbf{q}}$ is the probability density function for $\mathbf{q}$ (the bivariate normal density function with mean and covariance as provided in Table 1). Numerically approximating this integral using the 100×100 grid (a regular grid in $\mathbf{p}$-space but an irregular grid in $\mathbf{q}$-space), we have

$$\mathbb{P}(Y \in (Y_1, Y_2)) = \sum_{\mathbf{q}_i \,:\, y(\mathbf{q}_i) \in (Y_1, Y_2)} f_{\mathbf{q}}(\mathbf{q}_i) A_i,$$

where $\mathbf{q}_i$ is the i-th grid-point and A_i is the area in the grid associated with grid-point $\mathbf{q}_i$.

This approach is an alternative to Monte-Carlo sampling that is feasible since the parameter space is small. It may suffer from numerical approximation error, and if high accuracy in the probability density function for y was required, it would be important to investigate convergence of the computed density function with the number of grid-points, as well as confirming that the truncation of the parameter space to the square region used does not influence the results (note: the probability that $\mathbf{p}$ lies outside this space was numerically estimated to be less than 0.5% for all of the four computed distributions (corresponding to the four experimental datasets)). Since the probability distributions computed for Figure 6 are to illustrate uncertainty propagation only, and no strong quantitative conclusions were drawn, careful convergence analysis was deemed unnecessary for this case.

Maximum upstroke velocity

Figure S7 repeats Figure 6 using the maximum action potential amplitude V_{max} as the quantity of interest. The experimental results in this figure are taken from Cordeiro *et al.* 2008. It is interesting to note that the relationship between (V_0, K) and V_{max} is almost identical to that between (V_0, K) and upstroke velocity $\dot{V}_{\mathrm{max}}$ (compare Figure S7 (left) and Figure 6 (left)), and both quantities of interest are almost independent of K. This is the reason why the computed probability distributions in Figure S7 (right) for V_{max} have essentially the same shape as those shown in Figure 6 (right) for $\dot{V}_{\mathrm{max}}$.

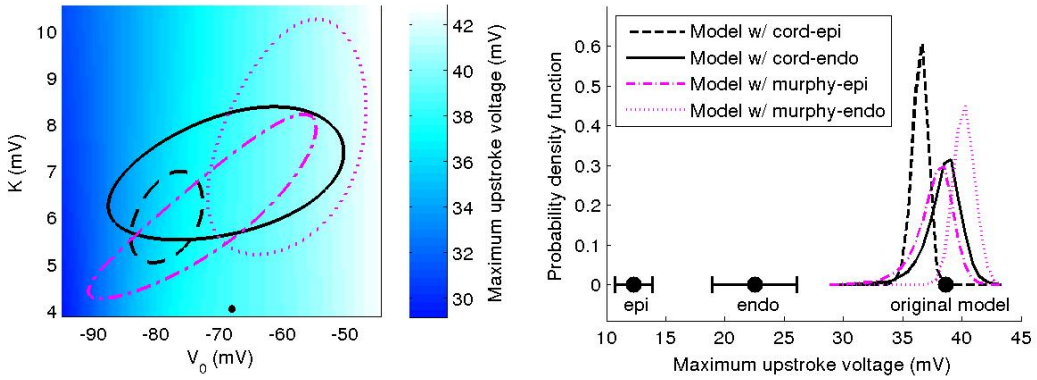

Figure S7: Propagation of uncertainty through a cardiac cell model using maximum action potential amplitude $V_{\max}$. Left: $V_{\max}$ as a function of the two parameters V_0 and K, together with NLME 95% prediction regions (for legend see right figure), and $(V_0^{\mathrm{F}}, K^{\mathrm{F}})$, the parameter choice used in the model (filled circle). Right: corresponding probability distributions for $V_{\max}$, as well as value computed using $(V_0^{\mathrm{F}}, K^{\mathrm{F}})$ and experimental measurements (filled circles).